



Universiteit
Leiden
The Netherlands

Indo-Slavic lexical isoglosses and the prehistoric dispersal of Indo-Iranian

Palmér, A.I.

Citation

Palmér, A. I. (2024, July 2). *Indo-Slavic lexical isoglosses and the prehistoric dispersal of Indo-Iranian*. Retrieved from <https://hdl.handle.net/1887/3765823>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3765823>

Note: To cite this publication please use the final published version (if applicable).

5. The archaeology and genetics of Indo-Iranian prehistory

5.1. Introduction

In section 4.7.7 above, it was concluded that lexical isoglosses shared by Indo-Iranian and Balto-Slavic provide evidence for a period of shared innovation that may be termed Indo-Slavic. As the Indo-Slavic period is intermediate between Core Proto-Indo-European and Proto-Indo-Iranian, it has implications for our understanding of the prehistoric dispersal of the Indo-Iranian languages. The aim of this chapter is to contextualize Indo-Iranian linguistic prehistory from archaeological and genetic perspectives. The focus, on the one hand, lies on the location of the Proto-Indo-Iranian homeland, and, on the other hand, on the dispersal of Pre-Proto-Indo-Iranian from the Proto-Indo-European homeland following the split of the protolanguage. For the latter question, three main scenarios will be presented and evaluated according to their compatibility with the linguistic evidence presented in Chapters 3 and 4. For reference, the most important archaeological cultures discussed throughout the chapter are summarized in Table 2.

	Period	Date (BCE)	Approx. location	Subsistence strategy
Yamnaya	EBA	3300–2600	Pontic-Caspian steppe	Mobile pastoralism (+ mixed farming west of the Dnipro)
Corded Ware	EBA–MBA	3000–2350	Northwest and Northeast Europe	Pastoralism, mixed farming
Fatyanovo-Balanovo (Corded Ware)	EBA–MBA	2900–2050	Northeast Europe, Dnipro to Vyatka-Kama interfluvium	Pastoralism, mixed farming
Bactria-Margiana archaeological complex	MBA–LBA	2250–1700	Central Asia, Amu Darya River	Irrigation farming
Abashevo	MBA	2200–1900	Middle Volga to South Urals	Sedentary pastoralism (+ mixed farming?)
Poltavka	MBA	2800–2100	Volga-Ural steppe	Mobile pastoralism
Sintashta	MBA	2100–1800	South Trans-Urals	Sedentary pastoralism
Alakul'-Fëdorovo	LBA	2000–900	Central Asian steppe	Mobile/Sedentary Pastoralism
Srubnaya	LBA	1850–1450	Eastern Pontic-Caspian steppe	Sedentary Pastoralism

Table 2. Summary of archaeological cultures discussed in the chapter. EBA = Early Bronze Age, MBA = Middle Bronze Age, LBA = Late Bronze Age.

5.2. The Indo-European homeland question

Although many homeland hypotheses have been proposed over the years (cf. Mallory 1989: 144), the debate on the Indo-European homeland has in recent decades been centred around the controversy between the Steppe hypothesis, the Anatolian hypothesis, and, albeit to a lesser extent, the Armenian hypothesis (see Gaitzsch & Tischler 2017).²⁰⁰

Proponents of the Steppe hypothesis (Benfey 1875; Tomaschek 1878: 862; Schrader 1890; Gimbutas 1956; Mallory 1989; Anthony 2007), which places the Indo-European homeland north of the Caucasus, between the Black Sea and the Caspian Sea, have relied on linguistic palaeontology as evidence for the connection between Proto-Indo-European culture and Early Bronze Age steppe cultures, termed Yamnaya (“pit grave”). A range of reconstructed terms, including words for wheeled vehicles and domesticated animals, delimit the timeframe of the Proto-Indo-European community to ca. 3500–2500 BCE, in which the Yamnaya culture (3300–2600 BCE, cf. Morgunova & Khokhlova 2013) provides

²⁰⁰ A notable alternative theory is Nichols’ (1997) “Bactria-Sogdiana” homeland, although she has now retracted this hypothesis. Incidentally, a Bactrian homeland was also proposed by Pictet (1859–1863).

a plausible origin for the dispersal of Indo-European to Europe and Asia, respectively (Anthony & Ringe 2015; Anthony 2023b).

Conversely, proponents of the Anatolian hypothesis (Renfrew 1987) have rejected arguments based on linguistic palaeontology as evidence in the homeland question. Their focus has instead lain on explaining the demographic processes behind the spread of the language family, arguing that the expansion of agriculture from Anatolia from ca. 7000 BCE provides a plausible vector for the spread of Indo-European (Bellwood 2001; 2013). Another argument comes from datings of Proto-Indo-European based on Bayesian phylogenetic analysis that are too early (ca. 8000–5000 BCE) to be compatible with the Steppe hypothesis (Gray & Atkinson 2003; Bouckaert et al. 2012; Heggarty et al. 2023). However, the early dating of Proto-Indo-European has largely been rejected by historical linguists, since it is incompatible with the evidence from linguistic palaeontology (Anthony & Ringe 2015; Kroonen et al. 2023). Moreover, the methodology is fundamentally based on the idea that rate of lexical replacement can be used to estimate divergence times of related languages (cf. Swadesh 1952), which is disputed (Bergsland & Vogt 1962; Nettle 1999).

Although the demographic argument was seen as a strong argument in favour of the Anatolian hypothesis, Allentoft et al. (2015) and Haak et al. (2015) have shown that (Indo-European-speaking) European and Central and South Asian populations have received significant gene flow from populations related to Pontic-Caspian steppe groups, forcing archaeologists to reconsider their views on the demographic dynamics between sedentary farmers and mobile pastoralists in prehistory. In fact, migrations of steppe populations caused massive population turnover in many parts of Europe and (to a lesser extent) Asia (Damgaard et al. 2018; Mathieson et al. 2018; Mittnik et al. 2018; Olalde et al. 2018; Narasimhan et al. 2019) on a scale that is compatible with the introduction and subsequent shift to a new language family. Taking the evidence from linguistic palaeontology and archaeogenomics together, the Steppe hypothesis comes out as the most plausible.

Lazaridis et al. (2022) argue that Proto-Indo-Anatolian may originate south of the Caucasus, with the non-Anatolian branches sharing a secondary homeland on the Pontic-Caspian steppe. This hybrid model in some way resembles the Armenian hypothesis (Gamkrelidze & Ivanov 1995), with the crucial difference that Indo-Iranian is still believed to have spread to Central and South Asia from the steppe region, rather than via the Iranian plateau. The advantage of the hybrid hypothesis is that it offers an explanation for the lack of steppe ancestry in Anatolia. However, the near-complete absence of reconstructable agricultural terms in Proto-Indo-Anatolian matches poorly with an Anatolian homeland, since this area was deeply agricultural (Kroonen et al. 2022). Thus, in this work, I place the Indo-European homeland in the 4th millennium Pontic-Caspian steppe.

Ultimately, for the purposes of this study, the difference between the Steppe hypothesis and the hybrid hypothesis of Lazaridis et al. (2022) is essentially inconsequential, since, in both models, the starting point of the Indo-Iranian dispersal (and the Balto-Slavic dispersal, for that matter) is the Early Bronze Age Pontic-Caspian steppe.

5.3. The Sintashta culture as an archaeological context for Proto-Indo-Iranian

The Sintashta culture encompasses around two dozen fortified settlements east of the Ural Mountains that share several material cultural and funerary features. The area is famous for the earliest attestation of the spoke-wheeled chariot in the late 21st century BCE (Lindner 2020). Besides the eponymous Sintashta site (Gening 1979), another major settlement was Arkaim (Kuz'mina 2007: 603). The culture is dated to 2100–1800 BCE (Anthony 2009: 57; Epimakhov, Zazovskaya & Alaeva 2023). It is thus chronologically intermediate between earlier Middle Bronze Age cultures west of the Urals such as Poltavka (2800–2100 BCE) and Abashevo (2200–1900 BCE) and Late Bronze Age cultures in Central Asia such as Alakul'-Fëdorovo²⁰¹ and Srubnaya (1850–1450 BCE).

The economy of the Sintashta culture was centred around pastoralism, as evidenced by the findings of domesticated animals of various species in burials. Judging from the proportion of bones found, the herd of Sintashta groups typically consisted of ~60 % cattle, ~25 % ovicaprids and ~15 % horse (Koryakova & Epimakhov 2007: 88). Single instances of pig or boar are also found, but pigs were not part of the herding economy (Zdanovich & Zdanovich 2002; Kuz'mina 2007: 146). Distinguishing sheep from goats is difficult without DNA analysis, but Kuz'mina (2007: 148) argues that sheep were more frequent than goats and that the latter are not found in burials as sacrificial animals, indicating that goats were less significant. Domesticated animals were a source for meat and milk (Zdanovich & Zdanovich 2002), which is confirmed by stable isotope analysis (Ventresca Miller et al. 2014; Hanks et al. 2018). Additionally, Judd et al. (2018: 11) argue that the lack of caries in individuals from Kamennyi Ambar-5 points to consumption of dairy products.

Although stockbreeding was the main subsistence strategy for the Sintashta population, Kuz'mina (2007: 141) argues that limited cereal farming was practiced as well. The proposed evidence for this consists of the placement of settlements, finds of stone querns, bronze sickles, and grain imprints on tools: all indirect evidence. Moreover, such tools may have been used for wild plants or other activities (Gerling 2015: 244; Mariotti Lippi et al. 2015). Zdanovich & Zdanovich (2002: 255) argue that the lands around Arkaim show traces of irrigation canals, pointing to earlier usage as fields for cultivation. Conversely, more recent studies stress the absence of any direct evidence for cereals in Sintashta settlements (Rühl, Herbig & Stobbe 2015; Judd et al. 2018). Absence of cereals is also supported by the lack of dental caries in Arkaim individuals (Anthony 2007: 405). Anthony mentions that charred millet grains found at Alandskoe have been taken as evidence for consumption of millet, at least at some sites, but widespread millet consumption in the Trans-Urals is not found during the Bronze Age, based on carbon and nitrogen isotope analysis (Ventresca Miller & Makarewicz 2019).

Previous research has drawn connections between Proto-Indo-Iranian and the Sintashta culture (Gening 1979), based on a combination of archaeological, linguistic, and genetic arguments.

²⁰¹ The terms Alakul' and Fëdorovo refer to what in earlier literature is known as the Andronovo culture, which is now regarded as inappropriate by many archaeologists (cf. Grigoriev 2021).

First, by applying a “retrospective approach”, Kuz'mina (2007: 163–64) derives the historically attested Iron Age steppe cultures of the Sauromatians and Saka peoples (who were Iranian-speaking) from the Sintashta culture. This argument is based on shared material cultural elements in these cultures, such as the strong equestrian tradition, similar types of arrows, spears, as well as other tools and weapons. Importantly, also non-functional elements of Sauromatian-Saka material culture have their roots in Sintashta culture, e.g., ceramic ornamentation, burial tradition, and aspects of the traditional dress, such as the pointy hat of the Saka. According to Kuz'mina (2007: 11), non-functional elements point to cultural identity.

Second, besides showing a cultural connection to historical Indo-Iranian-speaking communities, the chronology of the Sintashta culture roughly fits with an approximate dating of Proto-Indo-Iranian based on purely linguistic evidence. On the Indo-Aryan side, the relative chronology of the Vedas establishes the Rigveda (RV) as the oldest (cf. AiGr.), followed by the Atharvaveda (AV). Both clearly reference Panjab toponyms and were thus composed in South Asia (Witzel 1987). For the AV, a *terminus post quem* can be determined based on the mention of iron, which was widely used in South Asia from ca. 1000 BCE (Uesugi 2018: 4).²⁰² The earliest attestation of an Indo-Aryan language from the Mitanni kingdom can be dated to the 15th–14th centuries BCE (Witzel 1995: 99).²⁰³ On the Iranian side, the earliest direct attestation is represented by the Old Persian inscriptions from the 6th century BCE, next to Iranian personal names attested in Assyrian and Babylonian sources (Schmitt 1989: 25). The Avestan Gāthās, which reflect a linguistically more archaic stage than Old Persian, have been approximately dated to ca. 1000 BCE (Kellens 1989: 36). As in the case of the AV, the mention of iron in Avestan (cf. YAv. *hao-safnaēna-* ‘(made) of steel’, lit. ‘good iron’?) can be used to establish a *terminus post quem*. Although the exact geographical origin of the Avesta is unknown, the introduction of iron in the wider region of Iran and Central Asia begins ca. 1250–1000 BCE (Askarov 1999; Danti 2013). Together, the dating of the earliest Old Indo-Iranian texts implies a *terminus ante quem* for Proto-Indo-Iranian around 1500 BCE at the latest. Since the Rigvedic and Gāthic Avestan texts are so similar linguistically, the split cannot have been too long before this date. A split around ca. 2000 BCE fits well with the dating of the Sintashta culture to 2100–1800 BCE.

Third, a *terminus post quem* for Proto-Indo-Iranian can be approximated based on Indo-Iranian chariot terminology. Indo-Aryan and Iranian share a set of terms that can be reconstructed for Proto-Indo-Iranian, including **HratHa-* ‘chariot’, **HratHiH-* ‘chariot driver’, and **HratHai-štaH-* ‘chariot warrior’ (Malandra 1991; Oettinger 1994; Lubotsky 2023). Taken at face value, these words suggest that Proto-Indo-Iranian did not split before the invention of the spoke-wheeled chariot in the 21st century BCE. However, **HratHa-*

²⁰² The introduction of iron likely started already in the second millennium BCE, but was not widespread until after 1000 BCE. See further Uesugi (2018).

²⁰³ For an overview of the Mitanni Aryan language as an Indo-Aryan dialect, cf. Mayrhofer (1961). The Indo-Iranian presence in the Middle East may go back as early as the 18th century BCE, if *šāb ma-ri-ia-nim /šābīša ma-ri-a/ia-nim*, attested in the Leilan letter L.87–887 (cf. Eidem 2014: 142, fn. 16), reflects a Hurrian borrowing of Indo-Iranian **maria-* ‘young man, warrior’.

‘chariot’ is derived from an Indo-European word for (solid) wheel, **HrotHo-*, and could in principle have referred to a more primitive vehicle originally. On the other hand, the specific reference to the **HratHiH-* ‘chariot driver’ vs. **HratHai-štaH-* ‘chariot warrior’ strongly suggests a military context, implying a spoke-wheeled chariot; in the Near East, four-wheeled solid wheel wagons pulled by donkeys or onagers were used in military contexts as early as the 3rd millennium BCE, but there is no evidence for a similar practice in the steppe region (Hüttel 1994). Yet, it is striking that several Indo-Aryan technical terms relating to the spoked wheel, viz. Skt. *ará-* m. ‘spoke’, *nemí-* f. ‘wheel rim’, *paví-* m. ‘metal felly’, are not paralleled in Iranian. Although this is an *argumentum ex silentio*, it could be interpreted as evidence that the split of Indo-Iranian preceded the invention of the chariot (cf. Lubotsky 2023).

Fourth, locating the Proto-Indo-Iranian homeland close to the Ural Mountains is suggested by the many loanwords from early Indo-Iranian into Uralic languages (cf. Holopainen 2019). While many loanwords are from Proto-Iranian or later, there is also a Proto-Indo-Iranian and potentially a Pre-Proto-Indo-Iranian layer (see further 5.4 below), indicating a continuous presence of Indo-Iranian speakers in the Ural region.

Fifth, a layer of loanwords into Proto-Indo-Iranian have been argued to come from an unknown language of the Bactria-Margiana archaeological complex (BMAC) (Lubotsky 2001b). The BMAC civilization was at its peak around 2250–1700 BCE (Lyonnet & Dubova 2021: 32). Around its fortified settlements, the BMAC people practiced irrigation farming, cultivating wheat, barley, lentil, pea, grass pea, chickpea, grape, apple, and flax (Spengler et al. 2014). Domesticated animals include cattle, sheep, camels, pigs, and donkeys (Lyonnet & Dubova 2021: 23–24). Some of the proposed loanwords, e.g., **iauūiā-* ‘canal’, **Hustra-* ‘camel’, **kHara-* ‘donkey’, **kačiapa-* ‘tortoise’,²⁰⁴ can plausibly be connected to the BMAC, and suggest that Indo-Iranians came into contact with BMAC groups from the north rather than the south, as they should otherwise have been familiar with such concepts (Lubotsky 2001b: 307). Contact between BMAC agriculturalists and steppe pastoralists may further be evidenced by finds of Andronovo ceramics in BMAC contexts (Salvatori 2008: 64).²⁰⁵ Finds of cotton at the Sintashta culture settlement Kamennyi Ambar suggest contacts with Central or South Asian cultures (Shishlina, Koryakova & Orfinskaya 2022). The fact that some loanwords show irregular correspondences between Indo-Aryan and Iranian (Lubotsky 2001b; Palmér 2019) suggests that the contact with BMAC groups happened as Proto-Indo-Iranian was disintegrating, postdating the earliest Uralic contacts, which again supports a north to south movement of Indo-Iranian speakers.

Finally, population genomics suggests that steppe ancestry (i.e., ancestry related to Yamnaya steppe herders) spread to South Asia from Central Asian Middle Bronze Age groups around 2000–1500 BCE (Narasimhan et al. 2019: 7). Furthermore, from 2100–1700 BCE, outlier individuals from BMAC sites resemble Central Asian MBA groups

²⁰⁴ The Russian tortoise, *Testudo horsfieldii*, is native to the area of the BMAC (cf. Uetz et al. 2022).

²⁰⁵ However, these ceramics belong to the Tazabag'yab culture, which is no longer considered to be closely related to other so-called Andronovo cultures by some archaeologists (cf. Grigoriev 2021: 5).

(Narasimhan et al. 2019: 4). By comparing modern Iranian-speaking populations and ancient populations of Central Asia, Guarino-Vignon et al. (2022) show that there is genetic continuity from the Iron Age, and that the ancient populations can be modelled as a mix between local BMAC and incoming Central Asian steppe groups. With regards to South Asia (India in particular), a direct link to the Sintashta population is complicated by the fact that the Y-chromosome haplogroup overwhelmingly found in Sintashta is R-Z2124 (Narasimhan et al. 2019: S Table 1), whereas modern Indians with haplogroups related to R1a mostly have R-Y3+ (Underhill et al. 2015). Granted, both subclades are derived from R1a-Z93, but the formation of R-Z2124 and R-Y3 predates the formation of the Sintashta culture (Poznik et al. 2016). It is possible that an unsampled steppe population, autosomally similar to Sintashta, but with different Y-chromosome haplogroups, brought Indo-Iranian to India (cf. 5.4 below).

In sum, a diverse set of arguments support the Sintashta culture as a plausible archaeological proxy for early Indo-Iranians. However, that it would correspond one-to-one to the Proto-Indo-Iranian homeland, from which all subsequent Indo-Iranian languages originate, is doubtful, based on genetic evidence and the uncertainties regarding chariot terminology. As the following section will show, a slightly more complex scenario, involving the Abashevo culture, may be required to explain all the facts.

5.4. The Abashevo culture as an archaeological context for Pre-Proto-Indo-Iranian

When attempting to trace the origins of the Sintashta culture, archaeologists seem to agree on the importance of the Abashevo culture (Anthony 2009). Although previously believed to be older, the Abashevo culture is now radiocarbon dated to 2200–1900 BCE (Molodin, Epimakhov & Marčenko 2014; Mimoxod 2022), preceding the Sintashta culture by just over 100 years. Divided according to the location of sites, three variants are recognized: the Middle Volga, Don-Volga, and South Ural Abashevo culture (Mallory & Adams 1997: 1), the latter overlapping geographically with the Sintashta culture. Parpola (2022) has taken the Abashevo culture as an archaeological proxy for Pre-Proto-Indo-Iranian.

The Abashevo culture followed the kurgan burial custom of the Early Bronze Age Pontic-Caspian steppe cultures and shows evidence of a rich metallurgical tradition with copper and arsenic bronze weapons and tools (Kuz'mina 2021). The economy was mainly pastoralist, with a herd consisting of ~60–70 % cattle, ~10–20 % ovicaprids, and at most 15 % horses and domesticated pigs, respectively (Koryakova & Epimakhov 2007: 65). This is similar to the Sintashta culture (Kuz'mina 2007: 146), except for the inclusion of the domesticated pig. Parpola (2015: 55) has argued that metal sickles and stone querns provide evidence for agriculture, but Kuz'mina (2021) states that there is no direct evidence for farming. Pig husbandry is often taken as an indirect sign of agriculture, since they feed on rest products, but Koryakova & Epimakhov (2007: 65) argue that Abashevo pigs may have been fed acorns instead.

The argument that the Sintashta culture derives from the Abashevo culture is partly based on similarities in material culture. Not only have Abashevo pots been found in Sintashta burials, but the Abashevo ceramic tradition is argued to have influenced Sintashta pottery (Anthony 2007: 382; Koryakova & Epimakhov 2007: 74). Additionally, Sintashta weapons, tools, and adornments show influence from Abashevo precursors (Kuz'mina 2021). The connection between the cultures also makes sense from a geographical-chronological perspective, since the Abashevo culture spread eastward from the Middle Volga region, across the Urals, to the area of the Sintashta culture, shortly before the emergence of the latter (Anthony 2007: 382; Epimakhov 2020; Pärpola 2022: 15).

As for genetic evidence, Engovatova et al. (2023) present the first publication of samples from Abashevo individuals, all male ($n = 14$). Seven individuals carry Y-chromosome haplogroup R1a-Z93 and therefore show a plausible relationship to the Sintashta population.

Furthermore, linguistic evidence for language contact between Indo-Iranian and Uralic languages may support a connection between early Indo-Iranian speakers and the Abashevo culture. As mentioned in 5.3 above, there are loanwords in Uralic from Proto-Indo-Iranian and potentially even from Pre-Proto-Indo-Iranian (Holopainen 2019). The earliest loanwords were likely borrowed into an already dialectally differentiated post-Proto-Uralic stage (Common Uralic), as evidenced by their distribution in the western branches of Uralic, excluding Samoyed. Especially important are Common Uralic **mekši* 'honeybee' and **meti* 'honey', attested in Finnic, Mordvin, Permic, and Hungarian. Due to their vocalism, it has been argued that these words may have been borrowed from the Pre-Proto-Indo-Iranian ancestors of PIIr. **makši*- 'bee, fly', **mad^hu*- 'honey' < Pre-PIIr. **mekši* and **med^hu*-, respectively (Pärpola 2022: 17–18). Since apiculture was not practiced east of the Urals at the time, these loanwords likely entered Common Uralic as speakers migrated west from the Proto-Uralic homeland east of the Urals (Grünthal et al. 2022). Early Uralic speakers were likely associated with the westward spread of the Sejma-Turbino phenomenon (Zeng et al. 2023), dated to 2200–1900 BCE (Marchenko et al. 2017), which came into contact with the Abashevo culture (Černyx & Kuz'minyx 1987).

Although calling this layer of borrowings Pre-Proto-Indo-Iranian is consistent with linguistic reconstruction and archaeological facts, it must be borne in mind that vowel substitutions in Indo-Iranian-Uralic loanwords are notoriously difficult to interpret phonetically (Kümmel 2019). In the case of **makši*- 'bee, fly', since there are no cognates in other Indo-European languages (EWAia II: 287), the reconstruction of Pre-PIIr. **mekši*- rather than **mokši*- is based on Uralic, and therefore not decisive.²⁰⁶ In the case of **meti* 'honey', it is difficult to exclude that Uralic **e* reflects Proto-Indo-Iranian **a*.²⁰⁷ Thus,

²⁰⁶ The same is true for the idea that Common Uralic **ertä* 'side (of the body)' is borrowed from Pre-PIIr. **Herd^ho*- (Holopainen 2019: 81; Pärpola 2022: 18), ancestral to Skt. *árdha*- m. 'side, part, region', *ardhá*- m/n. '(one) half'. No Indo-European cognates confirm the reconstruction of an *e*-grade in the root; rather, the Indo-Iranian situation suggests a nomen actionis **Hord^ho*- 'separation, division' (cf. Lubotsky 1988b: 71, fn. 21).

²⁰⁷ It could be argued that Common Uralic **kekrä* 'circular thing' and **kečrā* 'spindle', corresponding to Skt. *cakrá*- m./n. 'wheel' and *cat(t)ra*- n. 'spindle', provide more convincing evidence that Uralic **e* reflects Pre-PIIr. **e*, since they appear to have been borrowed before the Proto-Indo-Iranian palatalization of **k^(h)* > **č*. However,

while the words plausibly link early contacts between Indo-Iranian and Uralic to the Abashevo-Sejma-Turbino context, determining the linguistic layer as specifically Pre-Proto-Indo-Iranian, as opposed to Proto-Indo-Iranian, must be considered uncertain.

In fact, the view that Sintashta and Abashevo reflect Proto-Indo-Iranian and Pre-Proto-Indo-Iranian, respectively (Parpola 2022), may be overly simplistic. The cultures overlap chronologically and geographically with each other, and even if there is a difference in material culture, this need not correlate one-to-one with the linguistic situation. From the perspective of genetics, as discussed in 5.3 above, the Sintashta population does not provide a perfect fit for Indo-Aryan-speaking groups in South Asia. Since the Abashevo population is, as of yet, much less thoroughly sampled, one might wonder if the missing R-Y3+ haplogroup males, required to explain the prevalence of this haplogroup in India, are hidden here.

To explore this idea further, let us consider an area where the Abashevo culture may be a better archaeolinguistic fit for Proto-Indo-Iranian than the Sintashta culture. Proto-Indo-Iranian inherited two words for ‘domesticated pig’ from Proto-Indo-European, **suH-* and **porko-*, as evidenced on the one hand by YAv. *hū-* m. ‘pig’, MiP Pahl. *hūg* ‘pig’, Oss. I *xoy* / D *xu* ‘pig’, and on the other by YAv. *parsa-* m. ‘pig(let)’, Khot. *pā’sa* ‘pig, hog’.

Skt. *sūkará-* m. ‘boar’ has traditionally been adduced, but the formation is obscure (**sūka-* + *-rá-*?). It is conspicuously similar in form to MiP Pahl. *hūkar(ag)* ‘porcupine’ (MacKenzie 1986: xxii), which could point to a PIr. **suHkara-* ‘swine; porcupine’ that is etymologically distinct from **suH-ka-* ‘pig’.²⁰⁸ Although usually translated as ‘wild boar’, a specific connotation to wild rather than domesticated pigs is not evident from the earliest attestations:

RV VII.55.4ab

tvám sūkarāśya dardrhi táva dardartu sūkaráh

‘Keep tearing at the boar; let the boar keep tearing at you’ (Jamison & Brereton 2014: 948).

ŚS XII.1.48c

varāhēṇa pṛthivī samvidānā sūkarāya ví jihīte mṛgāya

‘...the earth, in concord with the boar, opens itself to the wild hog’ (Whitney 1905: 669)

The phrase *sūkarāya mṛgāya* ‘to the wild *sūkará-*’ could imply that *sūkará-* on its own was semantically underspecified and could refer to either domesticated or wild pigs. This agrees with the Middle and Modern Indo-Aryan material, where the descendants of **sūkará-* vary in meaning, cf. Pā. *sūkara-* m. ‘pig’, Nep. *sūgar*, *sūgur* ‘domesticated pig’, Si. (*h*)*ūrā* ‘boar, wild pig’.

Kümmel (2019) argues that Proto-Indo-Iranian **č* may have been realized as a palatal stop at an early stage, which could have yielded Uralic **k*.

²⁰⁸ In view of the irregular correspondence with YAv. *sukurāna-* m. ‘porcupine’, MoP *sugur(na)* ‘id.’, Wan. *sugun/r* ‘id.’ < **sūkurna-* and Psht. *škuṇ* ‘porcupine’, Bal. *sūkān*, *sīnkur* ‘id.’ < **sūkurna-* (cf. Morgenstierne et al. 2003), a substrate origin is possible. In that case, the word for ‘porcupine’ may have been related to the word for ‘needle’ in the substrate language, borrowed as Skt. *sūcī-* f. ‘needle’, YAv. *sūkā-* f. ‘needle’ (Lubotsky 2001b), and subsequently folk-etymologically associated with the inherited word for ‘pig’, i.e., **suH(-ka)-*.

Based on the contrastive stem PIIr. **uarāj^ha-* ‘wild boar’ (Skt. *varāhā-* m. ‘wild boar’, YAv. *varāza-* m. ‘id.’),²⁰⁹ PIIr. **suH(-ka)-* and **parća-* likely referred to domesticated pigs, although perhaps not exclusively. As mentioned in 5.3 above, according to Kuz’mīna single instances of pig or boar bones have been found in Sintashta contexts, but in general she argues that “the complete absence of the pig make[s] up the characteristic feature [...] of Indo-Iranian stock-raising” (Kuz’mīna 2007: 158–59). Koryakova & Epimakhov (2007: 88) report no evidence of domesticated pig at Sintashta sites.²¹⁰ Mallory (1994) argues that this is consistent with the loss of the Indo-European pig words in most Indo-Iranian languages. This assessment is at odds with the reconstructed Proto-Indo-Iranian situation. Conversely, Abashevo sites offer clear evidence that the pig was part of the typical Abashevo herd. It is of course possible that the pig words could have been retained in an exclusively Sintashta-based Proto-Indo-Iranian community, through contact with nearby cultures that did keep domesticated pigs. However, together with the evidence for contact with Uralic, the pig words could be taken as evidence that part of the Proto-Indo-Iranian community should be identified with the Abashevo culture.

Expanding the Proto-Indo-Iranian homeland to include the Abashevo culture seems to be at odds with Proto-Indo-Iranian chariot terminology, however, since the Abashevans did not build chariots. Yet, as the discussion of the linguistic evidence for chariot technology has shown, most technical terms are not shared by Indo-Aryan and Iranian. The few terms that are shared (**HratHa-* ‘chariot’, **HratHiH-* ‘chariot driver’, **HratHai-štaH-* ‘chariot warrior’) are more general, and would also be compatible with a scenario where only parts of the Proto-Indo-Iranian community were building chariots, whereas the rest only knew of their existence (like the Abashevans likely did, given their cultural contact and proximity to the Sintashta culture). Once chariots had been invented, the technology quickly spread to the west of the Urals by the early 2nd millennium BCE (Kuznetsov 2006; Koryakova & Epimakhov 2007: 66; Kuznetsov & Mochalov 2016: 75), so the time gap between the formation of the Abashevo culture around 2200 BCE, the invention of the chariot 2050–2000 BCE, and the dissemination of the technology is rather insignificant.

Thus, two archaeolinguistic lines of evidence contradict each other in being consistent with either the Abashevo culture or the Sintashta culture as the Proto-Indo-Iranian homeland. As the discussion has shown, both the pig words and chariot terms may be explained away as valid linguistic palaeontological arguments by attributing their existence to cultural contacts rather than native cultural practices. Yet, there are other arguments linking both cultures to early Indo-Iranians, and since they are partly overlapping geographically, chronologically, and in terms of material culture, it is possible that the archaeological classification has little or nothing to do with the linguistic situation. Both cultures may represent parts of the Proto-Indo-Iranian homeland.

As argued in 4.4.1 above, Indo-Iranian attests some inherited agricultural terms, which indicate continuous familiarity with farming from Core Proto-Indo-European times.

²⁰⁹ The meaning ‘wild boar’ is supported by Skt. *varāhayú-* adj. ‘wishing for boar, boar-hunting’.

²¹⁰ Furthermore, the pig is not part of the herd in (presumably Indo-Iranian-speaking) Alakul’-Fëdorovo contexts, believed to derive from the Sintashta culture (Koryakova & Epimakhov 2007: 127; Kuz’mīna 2007).

This seems to be at odds with locating the Proto-Indo-Iranian homeland in the Sintashta culture, since there is ample evidence against agriculture being practiced by the Sintashta population (cf. 5.3 above). It is unclear if the hypothesis presented here, i.e., expanding the Proto-Indo-Iranian homeland to include the Abashevo culture, resolves this problem, since the presence of agriculture in the Abashevo culture is debated (cf. above). Even if neither the Abashevo culture nor the Sintashta culture practiced agriculture, it should be noted that these populations would have been in contact with the agriculturalists of the BMAC to the south (cf. 5.3 above), which could alternatively explain the presence of agricultural terms in Proto-Indo-Iranian.

5.5. From Yamnaya to Abashevo and Sintashta

As the previous sections have shown, the prehistory of the Indo-Iranian dispersal can be connected to the Sintashta and Abashevo cultures of the south Ural region in the end of the 3rd millennium BCE. For the preceding period, between the emergence of these cultures and the Indo-European homeland, there are several hypotheses that outline alternative scenarios for how speakers of Indo-Iranian reached the south Ural region.

5.5.1. Scenario 1: Eastward migration hypothesis

The most widely held hypothesis on how (Pre-Proto-)Indo-Iranian spread to the south Ural region is what I call the *Eastward migration hypothesis*. Individual variations aside, its proponents hold that the Proto-Indo-Iranian linguistic community of the south Urals was the result of a (north)eastward migration from the Indo-European steppe homeland during the 3rd millennium BCE. In a way, this may be thought of as the default hypothesis of Indo-Iranian origins, since a direct eastward migration is the shortest route from the steppe to the Ural region. That is not to say that the hypothesis is only based on geographical proximity, however.

Although Gimbutas (1963) connects Indo-Iranian to the Alakul'-Fëdorovo cultures, she makes no explicit mention of how the speakers got there from the Indo-European homeland. Mallory (1989: 263) follows Gimbutas' identification, and adds that the precursor of Indo-Iranian likely developed east of the Volga in the 3rd millennium BCE, corresponding to the Poltavka culture (Mallory & Adams 1997: 440). Also Parpola (2012; 2015; 2022; cf. also Carpelan & Parpola 2001) has explicitly connected Pre-Proto-Indo-Iranian to the Poltavka culture in his scenario of the prehistory of Indo-Iranian. Kuz'mina (2007: 305) agrees that Pre-Proto-Indo-Iranian may be connected to the Poltavka culture.

The Poltavka culture is essentially a Middle Bronze Age descendant of the Yamnaya culture that developed on the steppe between the Volga and Ural rivers ca. 2800–2100 BCE (Chernykh 1992: 132). It continues the kurgan burial tradition of the Yamnaya culture but is characterized by new ceramic styles and an increase in metallurgy. Unlike its contemporaneous western neighbour, the post-Yamnaya Catacomb culture, the copper used in Poltavka mainly came from the Ural region (Chernykh 1992: 133). The pastoralist economy was dominated by ovicaprids, supplemented by cattle and horses (Kuznetsov & Mochalov 2016: 86), and evidence for agriculture is lacking (for the lack of dental caries in

Poltavka individuals, cf. Murphy & Khokhlov 2016: 170–171). Like Yamnaya, the Poltavka culture herders were mobile, which makes an agricultural subsistence all the more unlikely (Anthony 2016: 3–6).

From an archaeological perspective, the Eastward migration scenario makes sense, as both the Abashevo and Sintashta cultures have been argued to show significant influence from the Poltavka culture (Anthony 2007: 383, 386; Parpola 2015: 297; Kuznetsov & Mochalov 2016: 85). The precursor of Indo-Iranian would then have developed in the eastern fringe of the Yamnaya culture, spreading further northeast during the Middle Bronze Age and reaching the Ural region toward the end of the 3rd millennium BCE, forming the Abashevo and Sintashta cultures.

From the perspective of genetics, however, continuity between the Poltavka culture and the Abashevo/Sintashta cultures is much less evident. Poltavka individuals cluster very close to the Yamnaya population (Mathieson et al. 2015; Narasimhan et al. 2019), indicating population continuity between the Early and Middle Bronze Age periods. However, they lack the Early European Farmer component found in Sintashta populations (Mathieson et al. 2015). Furthermore, Poltavka males generally carry Y-chromosome haplogroup R1b, associated with Yamnaya males, specifically the subclade R-Z2103 (Narasimhan et al. 2019: S Table 1). This haplogroup is also found in four samples from an Abashevo context (Engovatova et al. 2023), but is absent from Sintashta samples and later Central and South Asian populations associated with Indo-Iranian speakers.²¹¹ Based on this, the Poltavka population is implausible as a source for the Sintashta culture population and later groups related to the Indo-Iranian dispersal.

However, among the nine sampled individuals from Poltavka sites published by Mathieson et al. (2015) and Narasimhan et al. (2019), there is an outlier (sample I0432) that resembles Sintashta groups, showing admixture between steppe-related and European Farmer-related ancestry and carrying Y-chromosome haplogroup R1a-Z93 (specifically the Sintashta-like subtype R-Z2124, cf. Mathieson et al. 2015: S11). The individual is carbon dated to 2925–2536 calBCE and could provide a genetic link between the Poltavka and Sintashta cultures. According to Mathieson et al. (2015), the lack of additional evidence for this type of ancestry in Poltavka contexts could be explained by assuming that R1a males persisted in the area since the Chalcolithic, but were excluded from kurgan burials. However, in addition to being impossible to prove, this scenario was based on the fact that, at the time, males with Y-chromosome haplogroup R1a-Z93 had not been found elsewhere among ancient Europeans; this changed with Saag et al. (2021), who found ample evidence for such lineages in individuals from the Fatyanovo-Balanovo culture, an eastern extension of the European Corded Ware cultures. The latter population provides a more plausible source for later Central Asian groups such as Sintashta (cf. 5.5.2 below). Still, it is puzzling why a single individual matching the genetic signature of Sintashta groups would appear several hundred years prior to the formation of the Sintashta culture, genetically isolated

²¹¹ Since all but one of the hitherto sampled Abashevo individuals come from the Pepkino mass grave, presumably the result of a battle, it is conceivable that the buried individuals came from different cultural groups.

from the rest of the Poltavka samples.²¹² Importantly, there is no plausible nearby source for the European Farmer-related ancestry found in the Poltavka outlier, making its ancestry type difficult to explain as a local development in the context of the Poltavka culture.

Thus, while some details remain unclear, the overall impression is that the genetic evidence does not mirror the archaeological continuity between the Poltavka culture on the one hand and the Abashevo and Sintashta cultures on the other. This has consequences for the Eastward migration hypothesis that have not been acknowledged by its proponents. Most importantly, if the linguistic origins of Indo-Iranian lie in the context of the Poltavka culture, it requires the assumption of a language shift in the groups that would form the Sintashta culture, since the populations are so divergent genetically. It is difficult to imagine that such a language shift would have taken place without leaving traces in the Sintashta population or subsequent Indo-Iranian-speaking groups. Even in a scenario with language shift in the Sintashta population due to elite dominance of an Indo-Iranian-speaking minority with Poltavka origins, some genetic trace, if not in the autosomal DNA, then in Y-chromosome haplogroups, would be expected.

5.5.2. Scenario 2: via-Corded Ware hypothesis

Genetic evidence betrays a close relationship between the Sintashta population and Corded Ware groups of eastern Europe (Allentoft et al. 2015; Damgaard et al. 2018; Narasimhan et al. 2019). While Yamnaya groups can be modelled as a mix of Eastern Hunter Gatherer and Caucasus Hunter Gatherer ancestry (Allentoft et al. 2015; Haak et al. 2015), it has been argued that an additional $\sim 1/7$ Anatolian Farmer ancestry is required (Wang et al. 2019; Lazaridis et al. 2022). In contrast, the Corded Ware population has a larger proportion of Anatolian Farmer-like ancestry, as well as a small amount of Western Hunter Gatherer ancestry (Allentoft et al. 2015; Haak et al. 2015). The Corded Ware population is thought to result from admixture between steppe migrants and European Farmer populations in the late 4th millennium BCE (Papac et al. 2021; Ringbauer et al. 2024).²¹³ In the eastern Corded Ware populations belonging to the Fatyanovo culture, the Anatolian Farmer-like ancestry component makes up $\sim 33\%$ of the genetic ancestry, and all sampled male individuals carry Y-chromosomes of haplogroup R1a-Z93 (Saag et al. 2021). This is strikingly similar to the Sintashta population, which shows similar levels of Anatolian Farmer-like ancestry and the same predominance of Y-chromosome haplogroup R1a-Z93 among males.²¹⁴ R1a-Z93 males have now also been found in an Abashevo context (Engovatova et al. 2023). Since the earliest sampled Fatyanovo individuals are carbon dated hundreds of years before the

²¹² One possible explanation is that the carbon dating is wrong, and that the Poltavka outlier (I0432) in reality belongs to a later layer (after 2200 BCE). Apparently, the grave from which the individual was excavated was cut through by a later burial associated with the Middle Bronze Age Potapovka culture (Mathieson et al. 2015: S11).

²¹³ The origin of the Corded Ware genetic ancestry profile is a hotly debated topic. The fact that Corded Ware males carry Y-chromosomes of haplogroup R1a, which is unknown in Yamnaya males (where haplogroup R1b is predominant), suggests that the steppe ancestry component in Corded Ware individuals is not identical to that of Yamnaya populations. However, Ringbauer et al. (2024) have shown that Corded Ware individuals share IBD segments with Yamnaya individuals, which proves that they share ancestors only a few hundred years back.

²¹⁴ The same traits are found in modern South Asian populations, albeit with significant admixture with other ancestry groups and more Y-chromosome haplogroup variation among males (Narasimhan et al. 2019).

emergence of the Abashevo and Sintashta cultures (Saag et al. 2021), these can plausibly be explained as resulting from migrations of Fatyanovo groups.

The *via-Corded Ware hypothesis* can also be supported by archaeological evidence. Long before aDNA evidence had become available, archaeologists described connections between Corded Ware cultures and the Abashevo culture, seen as an off-shoot from the eastern Corded Ware cultures otherwise known as Fatyanovo-Balanovo (Gimbutas 1965: 605; Anthony 2007: 380ff; Kuz'mina 2007: 305; Nordqvist & Heyd 2020). As discussed in 5.3–5.4 above, the Sintashta culture is closely related to the Abashevo culture, and may therefore be considered to be indirectly related to the Corded Ware complex.

The Fatyanovo and Balanovo cultures make up the eastern part of the Corded Ware horizon, which stretches across the northern half of Europe from the Netherlands to the Volga, occupying the forest-steppe zone. The traditional view has been that the Fatyanovo culture formed as a result of western impulses from central Europe (possibly mediated via the Middle-Dniepr culture, cf. Anthony 2007: 380). Although the chronological difference is small, this view seems to be supported by recent radiocarbon dating, which gives 2900 BCE as an upper boundary for Fatyanovo (Saag et al. 2021), compared to the earliest finds of Corded Ware in Bohemia dating as far back as 3000 BCE (Papac et al. 2021). The latest dated Fatyanovo individual has a lower boundary of 2047 BCE (Saag et al. 2021), and there are charcoal remains dated between the 22nd and 18th centuries BCE, but in general most dates cluster around the early to middle 3rd millennium BCE (Nordqvist & Heyd 2020).

The Fatyanovo culture is mainly known from burials, which (unlike in the Abashevo and Sintashta cultures) are flat earth graves containing various grave goods, but only rarely metal objects (for an overview, cf. Nordqvist & Heyd 2020). Kurgan burials are found further east in Balanovo contexts, which possibly reflects influence from steppe cultures. The Balanovo culture is also characterized by the existence of settlements, which appear to be absent from Fatyanovo. Evidence for copper metallurgy is solid but not abundant, and is stronger in the area of the Balanovo culture closer to the Ural region. The subsistence strategy of Fatyanovo-Balanovo groups is debated. It seems clear that these Corded Ware groups were the first pastoralists in the forest-steppe zone of eastern Europe, with evidence for pigs, ovicaprids, cattle and horses. It has generally been assumed that Fatyanovo-Balanovo groups practiced agriculture, but there is little to no hard evidence for it, perhaps owing, at least partly, to the scanty attestation of settlements.

In linguistic terms, the *via-Corded Ware* scenario implies that Indo-Iranian would have formed in a linguistic community deriving from groups of Indo-European speakers who moved into central Europe at the turn of the 4th–3rd millennium BCE, forming the Corded Ware cultures. As these groups expanded to the northeast, forming the Fatyanovo-Balanovo cultures, the first specifically Indo-Iranian sound changes may have occurred toward the end of the Balanovo horizon, or in the context of the Abashevo culture (i.e., ca. 2300–2100 BCE). In this scenario, the Poltavka culture, which also influenced both Abashevo and Sintashta culturally, would not have been linguistically Indo-Iranian, but would rather reflect other Indo-European-speaking groups, who may eventually have assimilated linguistically to Indo-Iranian in the 2nd millennium BCE.

The Fatyanovo culture has been associated with Balto-Slavic, often specifically Baltic speakers (Gimbutas 1956: 163; Carpelan & Parpola 2001: 88; Anthony 2007: 380; Kuz'mina 2007: 305; Parpola 2022: 13). Given the results of Chapters 3–4, showing evidence for a period of Indo-Slavic shared innovation, it becomes possible to view the Fatyanovo culture as a plausible archaeological context for the Indo-Slavic linkage (cf. Narasimhan et al. 2019).

5.5.3. Scenario 3: Bell Beaker hypothesis

The origin of the Abashevo culture has played a crucial role in the Eastward migration and via-Corded Ware hypotheses, since it is seen as the immediate ancestor of the Sintashta culture. Most archaeologists have considered the Abashevo culture to contain elements derived from the Corded Ware cultures as well as the Poltavka culture. There is another hypothesis, however, which contends that the Abashevo culture arose following a migration of Bell Beaker people from central Europe.

Without completely rejecting the idea of influence from local predecessors, Mimoxod (2022) argues that the Middle Volga Abashevo culture is “fundamentally different from the previous substrate, which is represented by the Fatyanovo culture” (p. 122). He argues that the Abashevo burials with wooden coffins and kurgans surrounded by pillar fences find parallels in Moravian and other central European Bell Beaker sites, but not in any local cultures of eastern Europe. The Abashevo kurgan tradition is argued to be partly due to steppe influence, however. Unlike previous researchers, Mimoxod rejects any continuation of Fatyanovo ceramics in the Abashevo culture.

The formation of the Abashevo culture just after 2200 BCE coincides with the 4.2 ka BP climatic event, which was a period of global climate change causing increased wintertime precipitation in higher latitude areas and aridization in lower latitude areas (Mimoxod et al. 2022). These conditions pushed pastoralists in parts of Europe to seek winter pastures in areas such as the Pontic-Caspian steppe. Mimoxod et al. (2022) hypothesize that Bell Beaker groups from the Carpathian basin for this reason migrated to the Middle Volga region, forming the Abashevo culture. However, apart from being the closest area from which Bell Beakers could have migrated to the Middle Volga, there is no independent evidence that the migration would have come from the Carpathian region.

If the formation of the Abashevo culture was the result of a migration of Bell Beaker groups from central Europe, this should have left a signal in the genetic ancestry of the Abashevo population. However, seven of the 14 Abashevo samples published so far carry Y-chromosomes of haplogroup R1a-Z93 (Engovatova et al. 2023), which is rather associated with Corded Ware groups of the Fatyanovo culture (Saag et al. 2021), as well as later Sintashta groups and other populations linked to Indo-Iranian speakers (Narasimhan et al. 2019). Bell Beaker people from central Europe would be a poor fit as a source for these populations, since they tend to have lower proportions of steppe-like ancestry (~46 %) and higher proportions of Anatolian Farmer-like ancestry (~43 %) compared to Sintashta individuals (Olalde et al. 2018). It remains possible that it was an unsampled Bell Beaker group, whose ancestry profile more closely resembled that of Sintashta/Corded Ware

groups, that formed the Abashevo culture. However, for the time being, the migration assumed by Mimoxod (2022) cannot be considered supported by genetic evidence.

From a linguistic perspective, the *Bell Beaker hypothesis* would imply that Pre-Proto-Indo-Iranian was spoken somewhere in central Europe, perhaps in the Carpathian region, until just before 2200 BCE, from where it spread to the Middle Volga region. Unlike in the via-Corded Ware scenario, where Pre-Proto-Indo-Iranian is part, albeit on the eastern margins, of the Indo-Europeanization of central Europe, the Bell Beaker scenario places Indo-Iranian in central Europe proper until almost a millennium after the dissolution of Core Proto-Indo-European, in a cultural context that otherwise has mostly been associated with Celtic and Italic groups (e.g., Anthony 2007: 367).

5.6. Integration with linguistic evidence

Having presented three hypotheses on the origins of Indo-Iranian based on archaeological and genetic evidence, the aim of this section is to determine which scenario is the most consistent with the linguistic evidence.

The present study has shown that Indo-Iranian shares a substantial set of unique lexical isoglosses with Balto-Slavic, of which at least five are shared innovations. If the conclusion of Chapter 4 is accepted, an Indo-Slavic linkage must have existed somewhere in space and time between the split of Core Proto-Indo-European (before 3000 BCE) and Proto-Indo-Iranian (after 2200 BCE).

In the Eastward migration scenario, Indo-Iranian developed on the eastern fringe of the Indo-European homeland, associated with the Poltavka culture, and subsequently the Abashevo culture. Parpola (2022: 15) argues that isoglosses shared by Indo-Iranian and Balto-Slavic, such as the RUKI rule, resulted from language contact between (Pre-Proto-)Balto-Slavic-speaking Fatyanovo-Balanovo groups and (Pre-Proto-)Indo-Iranian-speaking Abashevo groups. However, there is no indication that the RUKI rule would have been a contact-induced phenomenon rather than an inherited development. As argued in Chapter 1, the RUKI rule may be an old sound change that failed to phonologize in other branches. Similarly, satemization cannot plausibly be explained as a contact-induced change in a Fatyanovo-Balanovo-Abashevo context, since it also includes Armenian and Albanian, which are unlikely to ever have been spoken in the Middle Volga region (cf. Thorsø 2023). As for the Indo-Slavic lexical isoglosses, there is no indication that they would have resulted from contact, since they predate all branch-specific sound changes. In any case, a scenario where the shared Indo-Slavic features, whether contact-induced or vertically transmitted, developed in the Middle Volga region requires all attested descendant languages to originate from there. This does not seem likely for Balto-Slavic, for which a more western homeland has been proposed (Gimbutas 1956: 163; Anthony 2007: 380; Kuz'mina 2007: 305).

Additionally, the Indo-Slavic lexical isoglosses include two probable agricultural terms (**dʰoH-neh₂-* ‘grains’ and **peH-ou-* ‘chaff’, cf. 4.4.1). This makes the Poltavka culture and the Middle Volga region problematic as a staging ground for the Indo-Slavic linkage, since there is no evidence for cereal cultivation in the steppe east of the Dnipro

during the Early to Middle Bronze Age (Rassamakin 1999: 152; Cunliffe 2015: 96; Kuznetsov & Mochalov 2016; Murphy & Khokhlov 2016). Similarly, the agricultural terms inherited from Core Proto-Indo-European in Proto-Balto-Slavic and Proto-Indo-Iranian, such as **h₂erh₃-* ‘to plough’ (cf. 4.4.1), suggest that both branches originate from the western part of the Indo-European homeland, west of the Dnipro (Kroonen et al. 2022).

In the via-Corded Ware scenario, the Indo-Slavic linkage may be correlated with the northeastward expansion of the Fatyanovo culture from western Ukraine, starting around 2900 BCE and reaching the Middle Volga region well before the end of the 3rd millennium BCE. This fits well with the chronological boundaries of Indo-Slavic (ca. 3000–2200 BCE) determined by the split of Core Proto-Indo-European and emergence of Proto-Indo-Iranian. Moreover, this scenario is compatible with the agricultural vocabulary of Indo-Iranian, in the sense that Indo-Iranian would ultimately originate in the agricultural western Ukraine, the proposed homeland of Core Indo-European (Kroonen et al. 2022). As for the Fatyanovo culture itself, direct evidence for cereal cultivation is lacking, but archaeologists tend to believe that agriculture played a role in its subsistence, based on indirect evidence (Nordqvist & Heyd 2020). Furthermore, it could be argued that familiarity with agriculture is implied by the fact that the Fatyanovo population shows substantial admixture with a European Farmer-like population (Saag et al. 2021).

Additionally, in the via-Corded Ware scenario, if satemization is taken as a shared innovation of the satem branches, this would have to have occurred in the late 4th or early 3rd millennium BCE, in a disintegrating Core Proto-Indo-European-speaking western Ukraine. This could be consistent with the dispersal of Armenian, which has been argued to originate in a western post-Yamnaya Catacomb culture context (Anthony 2007: 92; Thorsø 2023).

In the Bell Beaker scenario, Pre-Proto-Indo-Iranian speakers would have been situated in central Europe until a rapid migration displaced them to the Middle Volga region around 2200 BCE. In principle, this is compatible with the existence of an Indo-Slavic linkage; since the Bell Beaker phenomenon was likely multi-ethnic, not being correlated closely with a single genetic population type, it is possible that Indo-Slavic speakers carried Bell Beaker culture without showing significant linguistic affiliations to other Indo-European groups usually connected to the Bell Beaker phenomenon, such as Celtic and Italic (cf. Anthony 2007: 367). However, it is not the most attractive scenario. As for the Indo-Slavic and Indo-Iranian agricultural vocabulary, it is compatible with a central European context, as presupposed in the Bell Beaker scenario, since cereal cultivation is attested here (cf. Heyd, Husty & Kreiner 2004).

Aside from agricultural vocabulary, another linguistic palaeontological variable is represented by words for ‘pig’. As discussed in 5.4 above, Proto-Indo-Iranian inherited both **suH-* ‘pig’ and **porko-* ‘pig(let)’ from Core Proto-Indo-European, indicating familiarity with domesticated pigs. Interestingly, pig husbandry is not mentioned as a feature of the Poltavka culture (Cunliffe 2015: 96; Kuznetsov & Mochalov 2016), but is securely attested in the Fatyanovo culture (Nordqvist & Heyd 2020), which seems to favour the via-Corded Ware hypothesis. The Bell Beaker hypothesis is more difficult to evaluate from this perspective, since the exact location of the Pre-Proto-Indo-Iranian community in this

scenario is unclear, but pig husbandry was likely present in most of central Europe (Caliebe et al. 2017). It may of course be argued that the Poltavka population could have words for ‘pig’ without breeding them, in which case the evidence would not be incompatible with the Eastward migration scenario. However, the Proto-Indo-Iranian pig words are more consistent with the via-Corded Ware or Bell Beaker hypotheses.

A third line of evidence that may be indicative of the migration route of Pre-Proto-Indo-Iranian speakers is substrate words, i.e., words borrowed from non-Indo-European languages in prehistory. All branches of Core Indo-European in Europe, including Armenian, have been argued to share substrate words that may have been borrowed from pre-Indo-European languages of hunter-gatherer and farmer populations of Europe (cf. Schrijver 1997; Kroonen 2012; Jakob 2023a; Thorsø 2023; Wigman 2023 with lit.). A defining feature of most substrate words is formal irregularities that preclude a Proto-Indo-European origin. Many substrate words belong to semantic fields such as local flora and fauna, as well as agricultural terminology. Although they are in the minority, some substrate words are attested in branches whose historical locations are far apart, indicating that they were borrowed at a time when the branches were still located in closer proximity to each other.

If the Indo-Iranian branch originates from Indo-European populations that migrated to central or eastern Europe, before spreading east to the Ural region, as proposed in the via-Corded Ware and Bell Beaker hypotheses, we would expect to find traces of substrate words shared with European branches in Indo-Iranian languages. An exhaustive study is beyond the scope of this work, but a few potential cases may be discussed.²¹⁵ First, the Indo-Slavic isogloss **h₂eǵ-* ‘goat’ was argued to be a borrowing, with an irregular correspondence **h₂eiǵ-* ‘goat’ in Greek, Albanian, and Armenian (cf. 3.2.2). Although the meaning – seemingly belonging to a pastoralist semantic field – is not typical for a European substrate word, it represents a possible case linking Indo-Iranian to a European context. As for words with agricultural meaning, as discussed in 4.4.1, Iranian **H(a)rab^(h)anTa-* ‘chickpea’ and **(H)(a)uić-* ‘oats’ are possible comparanda of the European substrate words Gr. ἐρέβινθος m. ‘chickpea’ and PSl. **ovъsъ* m. ‘oats’, respectively. Especially **H(a)rab^(h)anTa-* ‘chickpea’ is difficult to reject, given the formal and semantic similarity to Greek. However, given its isolated attestation in a few Pamir languages and absence from Old Indo-Iranian languages, it is uncertain whether it goes back to Proto-Indo-Iranian. Another possible substrate word shared with a European branch is Skt. *kapāla-* n. ‘bowl, skull’ ~ OE *hafola* m. ‘skull’ < **kapolo-*, cf. also Lat. *caput* n. ‘head’ (EWAia I: 300). Given the required reconstruction of **a*, this is unlikely to be a native Indo-European word (Lubotsky 1989). However, since Skt. *kapāla-* structurally resembles substrate words from a later, post-Proto-Indo-Iranian stratum (Lubotsky 2001b),²¹⁶ it may be a much younger borrowing. In a similar semantic field, there is Skt. *kumbhá-* m. ‘jar, pitcher’, YAv. *xumba-* m. ‘pot’, which may be compared to Gr. κύμβη f.

²¹⁵ Indo-Iranian languages have been argued to reflect a Central Asian substrate, associated with the BMAC (Lubotsky 2001b; Witzel 2003).

²¹⁶ I.e., the “trisyllabic nouns with long middle syllable” (Lubotsky 2001b: 303).

‘cup, bowl’ (cf. EWAia I: 370). However, the Sanskrit and Avestan words do not match formally (**k-* vs. **kH-*),²¹⁷ suggesting that they were borrowed after the split of Proto-Indo-Iranian. In this case, they cannot be projected back to a European context.

Thus, there are some possible European substrate words in Indo-Iranian, although only **h₂eǵ-* ‘goat’ may be securely back-projected to Proto-Indo-Iranian, given its Indo-Slavic origin. In this sense, the Indo-Iranian situation is not entirely incompatible with the via-Corded Ware and Bell Beaker hypotheses. However, the substrate material appears much more limited in comparison to the European branches, which can be taken as an argument in favour of the Eastward migration hypothesis, where Pre-Proto-Indo-Iranian speakers remained far away from the pre-Indo-European linguistic landscape of Europe.

Yet, a caveat for this discussion is that it is unclear exactly when most substrate words were adopted into the European branches. The formation of the Corded Ware population, resulting from admixture between steppe populations and European Farmers (Papac et al. 2021; Ringbauer et al. 2024), is a plausible scenario for the adoption of some of the earliest substrate words. However, some substrate words show irregular correspondences within branches, e.g., PSl. **ovbs̃* m. ‘oats’ vs. Lith. *aviža* f. ‘oats’, suggesting that they reflect a later stratum of loanwords. Accordingly, the question is to what extent substrate words from the earliest stratum would show formal irregularities between the branches, or if such words would rather appear as regular Indo-European etyma. For example, Indo-Slavic **dʰoH-neh₂-* ‘grains’ has no compelling Indo-European etymology, and could reflect an early borrowing just like Indo-Slavic **h₂eǵ-* ‘goat’. More research is needed to clarify the origins of European substrate words, as well as to what extent such words are reflected in Indo-Iranian.

Following the above discussion, a summary of the compatibility of the three hypotheses on the Indo-Iranian dispersal with linguistic, archaeological, and genetic evidence is presented in Table 3.

²¹⁷ The Sanskrit and Avestan words can only be reconciled if one assumes that Grassmann’s Law affected voiceless aspirates, i.e., **kHumbʰa-* > **kʰumbʰa-*, or by assuming that Avestan underwent aspiration metathesis, i.e., **kHumba-* > **kʰumba-* > **kumbʰa-*.

	Linguistics				Archaeology	Genetics
	Indo-Slavic lexical isoglosses	Ilr. agricultural terminology	Ilr. pig words	Scarcity of Eur. substrate in Ilr.	Succession of archaeological cultures	Population continuity
Eastward migration	–	–	–	+	+	–
Via-Corded Ware	+	+	+	?	+	+
Bell Beaker	+	+	+	?/–	+	–

Table 3. Interdisciplinary compatibility of three hypotheses on Indo-Iranian origins.

To begin with, all three hypotheses are in principle compatible with the archaeological record, since the Abashevo and Sintashta cultures have been argued to be successors of the Poltavka culture, Fatyanovo culture, or Bell Beaker culture, respectively. However, when we incorporate population genomics into the picture, only the connection between the Fatyanovo culture and the Abashevo and Sintashta cultures correlates clearly with genetic evidence. With the caveat that the details surrounding the Poltavka outlier individual are still unclear (cf. 5.5.1), the Poltavka population does not seem to contribute to the genetic ancestry of Abashevo and Sintashta populations. This suggests that the influences of Poltavka material culture on Abashevo and Sintashta resulted from cultural contacts rather than migration. As for the Bell Beaker scenario, it does not seem compatible with the current genetic evidence.

Of the linguistic variables discussed, the Eastward migration hypothesis is inconsistent with Indo-Iranian agricultural terminology and pig words, since neither agriculture nor pig husbandry are features of the Poltavka culture. Similarly, the Poltavka culture does not provide a plausible context for the Indo-Slavic linkage, since the lexical isoglosses contain words with probable agricultural semantics. Conversely, the via-Corded Ware and Bell Beaker hypotheses seem consistent with the Indo-Slavic lexical isoglosses, as well as Indo-Iranian agricultural terminology and pig words. The one variable where the Eastward migration hypothesis has an edge over the other two is in the scarcity of European substrate words in Indo-Iranian. However, as discussed above, Indo-Iranian has a few potential European substrate words, which is why the compatibility of the via-Corded Ware hypothesis is marked with a question mark here. Perhaps the scarcity of European substrate words is most problematic for the Bell Beaker hypothesis, since in this scenario Pre-Proto-Indo-Iranian is argued to be situated in central Europe for hundreds of years following the split of Core Proto-Indo-European, whereas in the via-Corded Ware hypothesis it is associated with the Fatyanovo culture on the eastern fringes of Europe.

Thus, with regards to research question B (cf. 1.4), based on the results of the present study, the via-Corded Ware hypothesis is the most consistent with the combined evidence from linguistics, archaeology, and genetics. While the linguistic evidence previously adduced in favour of this scenario (i.e., satemization, RUKI rule, cf. Narasimhan et al. 2019) was determined to be ambiguous in Chapter 1, the Indo-Slavic lexical isoglosses present additional evidence in its favour, which, combined with linguistic palaeontological considerations, is consistent with an Indo-Slavic linkage that is correlated archaeologically and genetically with the Fatyanovo-Balanovo culture in the eastern Corded Ware horizon. A model of the prehistoric dispersal of Indo-Iranian based on this scenario is presented below (Figure 13).

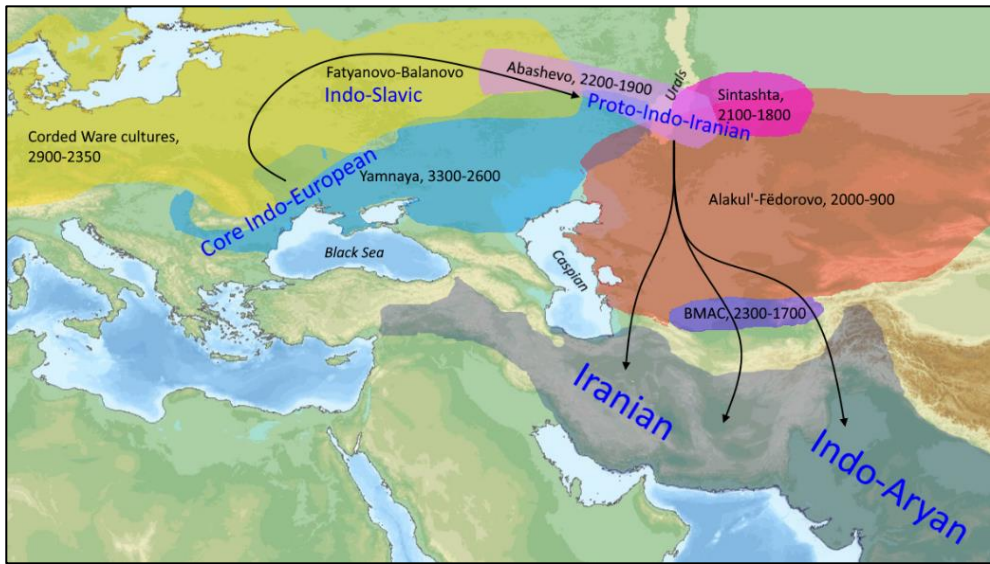


Figure 13. Model of the prehistoric dispersal of Indo-Iranian in the via-Corded Ware scenario. Archaeological cultures are given with dates BCE. Names for the chronological stages in the development from Core Indo-European to Indo-Iranian are indicated in blue. Approximate dispersal route of Indo-Slavic and Indo-Iranian is marked with arrows.

5.7. Limitations and outlook

The aim of this work has been to investigate the prehistoric dispersal of Indo-Iranian, specifically in the period between Core Proto-Indo-European and Proto-Indo-Iranian, by studying the phylogenetic relationship between Indo-Iranian and Balto-Slavic. The study of lexical isoglosses shared by these branches resulted in the postulation of a period of shared innovation that may be referred to as the Indo-Slavic linkage.

The discussion in 5.6 above has argued that the Indo-Slavic linkage is consistent with a scenario in which Indo-Iranian did not spread directly eastward from the Core Indo-European homeland, but rather moved to northeast Europe during the 3rd millennium BCE and gradually spread eastwards to the Ural region. However, as concluded in Chapter 4, the

phylogenetic position of Indo-Iranian and Balto-Slavic with respect to the other branches, particularly Greek and Germanic, is still to be determined. If Indo-Iranian and Balto-Slavic can be shown to also share innovations with other branches, this must be taken into account in hypotheses on their prehistoric dispersal. The possibility of Indo-Balto-Germanic shared innovations can probably be accounted for in the *via-Corded Ware* hypothesis, since Germanic has also been connected to the Corded Ware horizon (Anthony 2007: 360). On the other hand, Greek and Armenian have generally not been associated with the Corded Ware cultures, but rather with the Catacomb culture, developing out of the western Yamnaya horizon (Anthony 2007: 368; Clemente et al. 2021; Thorsø 2023). Therefore, if Indo-Iranian can be shown to share innovations with these branches to the exclusion of Balto-Slavic, this may have implications for the *via-Corded Ware* hypothesis.

The attempt to connect the Indo-Slavic and Proto-Indo-Iranian linguistic communities to archaeological contexts has utilized the methodology known as linguistic palaeontology. In some cases, it proved difficult to find linguistic material that could disambiguate between archaeological cultures. For example, Indo-Slavic was argued to have five unique terms related to dairy production (cf. 4.4.2), but since all relevant archaeological cultures (Yamnaya, Poltavka, Fatyanovo, Abashevo, Sintashta) likely used dairy products as part of their subsistence, the linguistic evidence is not very informative. In other cases, however, the linguistic material was able to provide important insights when compared to the archaeological record. Words relating to chariots and apiculture allowed the Proto-Indo-Iranian homeland to be correlated with the Abashevo and Sintashta cultures. In the Pre-Proto-Indo-Iranian period, agricultural terms and pig words point in favour of the Fatyanovo culture over the Poltavka culture. Yet, as the discussion of these semantic fields has shown, the material is often compatible with conflicting interpretations, due to the limitations of semantic reconstruction. In particular, the distinction between wild and domesticated cereals and animals is often difficult to establish with a high degree of certainty. In these cases, it is rather a matter of determining the most probable interpretation, and comparing this to other lines of evidence (from linguistics, as well as archaeology and genetics), keeping in mind that new material may appear in the future that strengthens or weakens the chosen interpretation.

This thesis has explored how linguistic considerations relating to phylogenetic subgrouping and linguistic palaeontology can be correlated to archaeological and genetic evidence, in order to reconstruct the prehistoric dispersal of the Indo-Iranian branch. It is hoped that future research will be able to fill the remaining gaps regarding Indo-European phylogeny that limit the conclusions of this study, as well as further refine the reconstruction of Eurasian population genomics, in order to reach a more complete understanding of Indo-European prehistory.